Solid-State Electronics Vol. 25, No. 8, pp. 741–747, 1982
Printed in Great Britain.

0038-1101/82/080741-07$03.00/0
Pergamon Press Ltd.

A PHYSICAL MODEL FOR THE DEPENDENCE OF CARRIER LIFETIME ON DOPING DENSITY IN NONDEGENERATE SILICON†

J. G. FOSSUM and D. S. LEE

Department of Electrical Engineering, University of Florida, Gainesville, FL 32611, U.S.A.

(*Received* 12 *July* 1981; *in revised form* 3 *December* 1981)

Abstract—A theoretical model that describes the dependence of carrier lifetime on doping density, which is based on the equilibrium solubility of a single defect in nondegenerately doped silicon, is developed. The model predictions are consistent with the *longest* measured hole and electron lifetimes reported for n-type and p-type silicon, and hence imply a possibly "fundamental" (unavoidable) defect in silicon. The defect is acceptor-type and is more soluble in n-type than in p-type silicon, which suggests a longer fundamental limit for electron lifetime than for hole lifetime at a given nondegenerate doping density. The prevalent, minimum density of the defect, which defines these limits, occurs at the processing temperature below which the defect is virtually immobile in the silicon lattice. The analysis reveals that this temperature is in the range 300–400°C, and thus emphasizes, when related also to common non-fundamental defects, the significance of low-temperature processing in the fabrication of silicon devices requiring long or well-controlled carrier lifetimes.

NOTATION

c_n^0	electron capture parameter of (neutral) defect
c_p^-	hole capture parameter of (ionized) defect
D^e	unionized defect in external phase
D^s	unionized defect in semiconductor (silicon) phase
D^-	ionized (acceptor-type) defect in semiconductor (silicon) phase
E_a	activation energy for defect formation in silicon
E_d	energy level (in energy gap) of defect state
E_i	intrinsic Fermi level
g_d	degeneracy factor of defect energy state
h^+	hole in semiconductor (silicon) phase
k	Boltzmann's constant
K_1, K_2, K_3	mass-action reaction (tempearture-dependent) constants
N_A	acceptor doping density
N_D	donor doping density
N_D^+	ionized donor density
N_d	defect density in semiconductor (silicon)
N_d^e	(constant) defect density in external phase
N_d^0	unionized defect density in semiconductor (silicon)
N_d^-	ionized (acceptor-type) defect density in semiconductor (silicon)
N_{d0}^-	ionized (acceptor-type) defect density in undoped semiconductor (silicon)
N_{Si}	atomic density of silicon
N	electron density
n_i	intrinsic carrier density
P	hole density
T	absolute temperature
T_f	effective temperature of defect formation in semiconductor (silicon)
U	net electron-hole recombination rate
τ_H	high-injection ($P \simeq N$) carrier lifetime
τ_n	minority electron lifetime
τ_p	minority hole lifetime

†This work was supported by Sandia National Laboratories under Contract No. 23-1161 and by United States-Spain Cooperative Research Grant No. T3773008.

INTRODUCTION

The performance of virtually every semiconductor device or integrated circuit is dependent on the carrier lifetimes in the semiconductor. For example, the power-conversion efficiency of solar cells and the quantum efficiency of detectors depend strongly on recombination carrier lifetimes, and the signal-to-noise ratio of dynamic MOS memories in VLSI circuits and the charge-transfer efficiency of CCDs depend critically on generation carrier lifetimes. Because silicon is the most commonly used semiconductor in the electronics industry, an understanding of how carrier lifetimes are related to the silicon growth and processing has practical significance.

The recombination and generation carrier lifetimes are defined by the predominant carrier recombination-generation mechanisms[1], most commonly the band-band (phonon-assisted) Auger-impact process and the capture-emission (Shockley-Read-Hall (SRH)[2]) process at defects. The latter process involves bound states, or traps, in the energy gap resulting from the defects in the semiconductor lattice. For nondegenerately doped silicon, the SRH process is generally dominant, and consequently the carrier lifetimes are controlled by the densities of defects in the silicon. To achieve optimal designs of silicon devices and circuits, we must know the upper bounds, or fundamental limits, for carrier lifetimes, and we must understand how the silicon processing influences the solubility of defects that define these bounds.

Many previous publications[3] have disclosed measured dependencies of carrier lifetime on doping density in silicon. However, because most of these observed dependencies resulted from "non-fundamental" defects, e.g. impurities, and because the data were obtained by a variety of experimental techniques, all of

which are subject to possibly significant error[3], the lifetimes are widely scattered. Thus inferences from these measurements are not generally significant and are often misleading.

To provide a better understanding of lifetime-vs-doping dependencies in silicon, we develop a theoretical model based on the presence of a single defect in nondegenerate silicon. The model derives from the analogy to chemical reactions of the dopant-induced formation of lattice defects[4, 5]. The determination of the defect (solute) density follows from a characterization, based on the law of mass action, of the doped semiconductor (solvent) in thermal equilibrium at the prevalent temperature of defect formation T_f. Derived expressions for the dependence of the defect density (assumed to be homogeneous) on the doping density in both n-type and p-type silicon, i.e. $N_d(N_D)$ and $N_d(N_A)$, which depend on T_f, enable the determination of the desired dependencies of minority hole and electron lifetimes, $\tau_p(N_D)$ and $\tau_n(N_A)$.

To support the model and to assay important defects, we compare theoretical predictions with measurements of recombination lifetime-versus-doping dependencies. By focusing our attention on those dependencies showing the *longest* carrier lifetimes ($\rightarrow$ 1 msec) reported for silicon, we detect a possibly "fundamental" defect in silicon. The defect, which is acceptor-type, appears to be fundamental (unavoidable) because (a) it is readily available in unlimited quantity for incorporation into the silicon lattice, and (b) its solubility is fundamentally related to the doping density and, although its density can be influenced by the processing, it cannot be totally eliminated from the lattice. The defect could therefore be a vacancy, but because singular vacancies are unstable in silicon at room temperature[6], it is more likely a divacancy or a vacancy complex[7]. Our analysis is not conclusive in this regard, but the results are curious and should induce additional work.

The analysis suggests that the defect is more soluble in n-type than in p-type silicon, and hence that the fundamental limit for electron lifetime is longer than that for hole lifetime at a given doping density. The prevalent, minimum density of the defect, which defines these limits, is "frozen-in" at the temperature T_f below which the defect is virtually immobile in the silicon lattice. Comparisons of our model predictions with ultimate carrier lifetime data reveal that T_f is in the range 300–400°C, and thereby emphasize the significance of low-temperature processing in the fabrication of silicon devices requiring long or well-controlled carrier lifetimes. For example, the efficacy of optimal annealing schedules is stressed, and furthermore a possible physical mechanism underlying such annealing is described.

Finally, the basis of our analysis, which is concentrated on a possibly fundamental defect, is used to qualitatively discuss common carrier lifetimes in silicon[3] that are defined by non-fundamental defects.

ANALYSIS

Consider first n-type silicon. Kendall[8] has reported measurements of $\tau_p(N_D)$ made on silicon wafers that had never been heated above 450°C. Although it was not stated whether or not this silicon was float-zone, we surmise that it was because the lifetimes are much longer than those found in crucible-grown silicon[3]. These measurements, which involved doping densities ranging from about $10^{14}\,cm^{-3}$ to about $10^{18}\,cm^{-3}$, yielded, to our knowledge, the longest recombination hole lifetimes ever reported for silicon. The data are described well by the empirical expression[9]

$$\tau_p(N_D) \simeq \frac{\tau_{0p}}{1+\dfrac{N_D}{N_{0D}}} \tag{1}$$

where $\tau_{0p} \simeq 4.0\times10^{-4}$ sec and $N_{0D} \simeq 7.1\times10^{15}\,cm^{-3}$; (1) shows that τ_p varies inversely with N_D when the resistivity is sufficiently low. The high-resistivity lifetime implied by (1) is consistent with the long hole lifetimes measured by Graff and Pieper[10].

The $\tau_p(N_D)$ dependence in (1) is also consistent with results of determinations[11], based on measured electrical characteristics, of recombination hole lifetime in the base regions of superior p^+nn^+ float-zone silicon solar cells fabricated at Sandia Laboratories. The consistency of these data and of those of Graff and Pieper[10], taken from carefully processed devices, with those reported by Kendall[8], taken from unprocessed material, portends the predominance of a "fundamental" (unavoidable) defect whose prevalent, minimum solubility in silicon results readily from proper processing.

The asymptotic behavior of (1) is characteristic of a singularly ionized, acceptor-type defect in n-type silicon. We now demonstrate this by deriving a theoretical model for the dependence $N_d(N_D)$ of the solubility of the defect on the doping density. We then use the model and the carrier lifetime data to estimate the physical properties of this possibly fundamental defect. This defect characterization also involves our derivation of the solubility $N_d(N_A)$ of the same defect in p-type silicon, which leads to $\tau_n(N_A)$ predictions that conform crudely with measurements[12, 13] of ultimate recombination electron lifetimes in silicon.

We emphasize that this analysis is based on the assumption of *homogeneous* distributions of defects in the silicon. Consequently it is applicable only to regions whose volumes are much larger than $1/N_d$ such that *macroscopic* (volume-average) carrier lifetimes faithfully characterize the recombination and generation of hole-electron pairs.

Consider the defect as a solute that is in (chemical) equilibrium with both a nondegenerately doped semiconductor (the solvent) and an external phase with constant activity, which implies an unlimited source for the (fundamental) defect[5]:

$$D^e \overset{K_1}{\rightleftarrows} D^s \overset{K_2}{\rightleftarrows} D^- + h^+. \tag{2}$$

In the reversible reactions indicated in (2), D^e and D^s represent the unionized defect in the external and semiconductor phases, and D^- and h^+ represent the ionized

defect and a hole in the semiconductor. This approach, which is equivalent to that based on the Shockley-Last theory[4], is used here because of the physical insight it provides and because of the simplifying assumptions it reveals; these attributes will become apparent as the analysis is described.

The law of mass action applied to (2) gives[5]

$$N_d^- P = K_2 N_d^0 = K_2 K_1 N_d^e \equiv K_3 \tag{3}$$

where N_d^- and N_d^0 are the ionized and unionized defect densities in the semiconductor, N_d^e is the constant defect density in the external phase, and K_1, K_2, and K_3 are mass-action (temperature-dependent) constants. Combining (3) with the hole-electron mass-action law for the nondegenerate semiconductor,

$$PN = n_i^2 \tag{4}$$

where P and N are the hole and electron densities and n_i is the intrinsic carrier density, and with the condition for electrical neutrality in the semiconductor doped with N_D donor impurities,

$$P - N + N_D^+ - N_d^- = 0, \tag{5}$$

yields the following expression of the equilibrium solubility of the ionized defect in the semiconductor:

$$N_d^- \simeq N_{d0}^- \left\{ \frac{N_D}{2n_i} + \left[1 + \left(\frac{N_D}{2n_i} \right)^2 \right]^{1/2} \right\} \tag{6}$$

where N_{d0}^- is the solubility of the ionized defect in the undoped semiconductor. In writing (6) we have recognized the common conditions that all the donor impurities are ionized, i.e. $N_D \simeq N_D^+$, and that $N_d^- \ll N_D$ in extrinsic (n-type) silicon.

Note that (6) is applicable at temperatures T high enough that the defect is sufficiently mobile in the silicon to support the equilibrium condition (2) and to allow N_d^- and N_d^0 to vary in accordance with (3). During the silicon growth or processing, when T finally drops below some temperature T_f, the defect becomes virtually immobile, and its density is hence "frozen" in accordance with (6) evaluated at $T = T_f$; i.e. $n_i = n_i(T = T_f)$ and $N_{d0}^- = N_{d0}^-(T = T_f)$. The temperature T_f, which we estimate later, is thus the effective temperature of formation of the defect and is fundamentally related to the mobility of the defect in silicon. Note that if the silicon is cooled too quickly (quenched) during the growth or processing, the defects may freeze-in at a temperature higher than T_f. The resulting defect density is then higher than the "fundamental limit" corresponding to (6) evaluated at $T = T_f$.

The total defect density N_d that is frozen-in at $T = T_f$, and hence affects τ_p at room temperature, is the sum of N_d^- given by (6) and N_d^0. We must therefore determine N_d^0 at $T = T_f$, which, as indicated by (3), is independent of N_D. The general relationship between N_d^0 and N_d^-, i.e. the degree of ionization of the defect, is quite complex, even if we assume a single energy level for the defect. It involves the degeneracy factor g_d of the defect energy state, which is unknown, and its energy level E_d, whose value relative to the intrinsic Fermi level E_i depends on T_f since both E_d and E_i vary with temperature.

Because of the uncertainty in the energy-band structure at $T = T_f$, we can only estimate the relationship between N_d^0 and N_d^-, which we do as follows. For the particular case $N_D \ll n_i(T_f)$, the hole and electron densities at $T = T_f$ are nearly equal, and for simplicity we assume that approximately half of the defects are ionized; this would be true, for example, if $g_d \simeq 1$ and $E_d \simeq E_i$ at $T = T_f$. Thus for this particular case, $N_d^0 \simeq N_d^-(N_D \ll n_i) \simeq N_{d0}^-$, as implied by (6). Since N_d^0 and N_{d0}^- are independent of N_D, their relationship, subject to the stated assumption, holds generally *for all* N_D. Hence using (6), we can write the explicit dependence of $N_d (= N_d^- + N_d^0)$ on N_D as

$$N_d(N_D) \simeq N_d^0 \left\{ 1 + \frac{N_D}{2n_i} + \left[1 + \left(\frac{N_D}{2n_i} \right)^2 \right]^{1/2} \right\} \tag{7}$$

where N_d^0 and n_i are evaluated at $T = T_f$. This model implies carrier lifetimes that are self-consistent with the hole lifetime data referred to earlier, as we will show, and its underlying assumptions are no more uncertain than the properties of the defect at $T = T_f$. Not only may these properties differ from those at room temperature, but the nature of the defect may be entirely different[14]; for example, silicon vacancies, which may be stable at $T = T_f$[6], might precipitate to form divacancies as T drops to room temperature.

A similar treatment of the density of the same defect in p-type silicon yields

$$N_d(N_A) \simeq N_d^0 \left\{ 1 - \frac{N_A}{2n_i} + \left[1 + \left(\frac{N_A}{2n_i} \right)^2 \right]^{1/2} \right\}. \tag{8}$$

We have checked the generality of (7) and (8) by examining the effects of assuming $N_d^0 = f N_d^-$ where $f \neq 1$. For $10^{-2} \leq f \leq 10^2$, we have found that the lifetime-vs-doping dependencies implied by (7) and (8) do not change significantly in the doping range where the SRH process predominates. Thus subject to the uncertainties associated with the defect, (7) and (8) seem to be a reasonable model.

In (7) and (8), N_d^0, which depends only on T_f, is the solubility of the neutral defect in silicon at the effective temperature of defect formation. Minimization of the lattice free energy at $T = T_f$ with respect to N_d^0 yields[15]

$$N_d^0(T_f) = N_{Si} \exp\left(\frac{-E_a}{kT_f} \right) \tag{9}$$

where N_{Si} is the atomic density of silicon ($\simeq 5 \times 10^{22}\,\mathrm{cm}^{-3}$) and E_a is the activation energy required to form the defect in the silicon lattice.

The dependencies on doping density given by (7) and (8) are illustrated in Fig. 1. For low doping densities [$N_D, N_A \ll n_i(T_f)$], both (7) and (8) yield $N_d \simeq 2N_d^0(T_f)$.

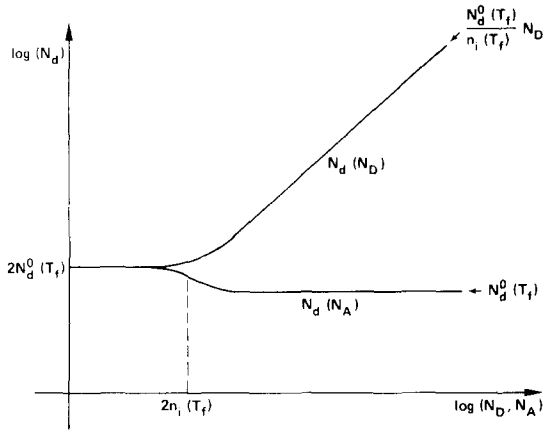

Fig. 1. Acceptor-type defect density dependencies on doping density for n-type [$N_d(N_D)$] and p-type [$N_d(N_A)$] nondegenerate silicon. The asymptotic behaviors are indicated.

For high doping densities [$N_D, N_A \gg n_i(T_f)$], the density of the defect in n-type material is dramatically higher than that in p-type material, and increases linearly with N_D: $N_d \simeq [N_d^0(T_f)/n_i(T_f)]N_D$. This asymptotic behavior is consistent with (1) and the $\tau_p(N_D)$ data discussed earlier. In p-type material, $N_d(N_A)$ is nearly insensitive to N_A, but actually decreases slightly with increasing N_A: $N_d \to N_d^0(T_f)$. This prediction is crudely consistent with $\tau_n(N_A)$ data published by Fisher and Pschunder[12] and by Graff and Fischer[13]. These data, which were taken on as-grown, float-zone silicon, show considerable scatter; however the *ultimate* electron lifetime increases, but by a factor of only two or three, as N_A increases to about $10^{17}\,\mathrm{cm}^{-3}$, above which τ_n decreases sharply because of Auger recombination.

To transform (7) and (8) to carrier lifetime dependencies on doping, which can be compared to measurements of $\tau_p(N_D)$ and $\tau_n(N_A)$, we use the SRH model[2] for carrier recombination through a single-level trap, including the effect of a non-unity degeneracy factor:

$$U = \frac{PN - n_i^2}{\left\{\frac{1}{c_p^- N_d}\left[N + \frac{1}{g_d} n_i \exp\left(\frac{E_d - E_i}{kT}\right)\right] + \frac{1}{c_n^0 N_d}\left[P + g_d n_i \exp\left(\frac{E_i - E_d}{kT}\right)\right]\right\}}. \tag{10}$$

The superscripts on the carrier capture parameters c_p^- and c_n^0 indicate the ionization state of the defect corresponding to the particular capture process. Since (10) will be used in conjunction with room-temperature data, n_i, E_d, and E_i in (10) represent room-temperature values.

For n-type material, the low-injection ($N \simeq N_D \gg P$) minority hole lifetime is defined as

$$\tau_p \triangleq \frac{P - \frac{n_i^2}{N_D}}{U}. \tag{11}$$

For p-type material ($P \simeq N_A \gg N$), the minority electron lifetime is

$$\tau_n \triangleq \frac{N - \frac{n_i^2}{N_A}}{U}. \tag{12}$$

For high-injection conditions ($P \simeq N \gg N_D, N_A$), the carrier lifetime is

$$\tau_H \triangleq \frac{P}{U} \simeq \frac{1}{N_d}\left(\frac{1}{c_p^-} + \frac{1}{c_n^0}\right). \tag{13}$$

The lifetime expressions given by (11)–(13), together with (10)—which generally can be simplified depending on the doping and the injection level in the semiconductor—and with (7) for n-type and (8) for p-type material, provide a theoretical basis for observed dependencies of carrier lifetimes on doping densities in silicon. For example, (7), (10), and (11) imply a $\tau_p(N_D)$ dependence that is consistent with (1), whereas (8), (10), and (12) yield only a weak $\tau_n(N_A)$ dependence in agreement with [12] and [13]. With (7)–(10) characterized to represent the fundamental defect that we have suggested, (11)–(13) can be taken as estimates of fundamental-limit carrier lifetimes in nondegenerately doped silicon.

To explicitly detail these estimates, we appraise properties of the defect by comparing quantitatively the predictions of (7)–(13) with the lifetime data. In doing this we are implicitly assuming that not only is the same (fundamental) defect predominant in all the silicon samples used in the independent experiments, but also that T_f is the same. This assumption is not unreasonable. We are considering only the longest carrier lifetimes reported for silicon, which prompts us to assume a fundamental defect, and, as we discussed earlier, T_f is fundamentally defined by the mobility of the defect in silicon, and is therefore not strongly dependent on the dopant type nor the doping density. Nevertheless, because of this assumption and because of the uncertainties associated with the defect, the following appraisal is not rigorous, but is stimulating.

Concentrating first on measured ultimate hole [8], [10], [11] and electron [12], [13] lifetimes in high-resistivity silicon [$N_d \simeq 2N_d^0(T_f)$], including the high-injection lifetime[11], we obtain

$$\frac{1}{c_p^-(2N_d^0)} \simeq 3 \times 10^{-4}\ \mathrm{sec}, \tag{14}$$

$$\frac{1}{c_n^0(2N_d^0)} \simeq 5 \times 10^{-4}\ \mathrm{sec}, \tag{15}$$

and, *if* $g_d \simeq 1$,

$$[E_d - E_i]_{T=300\,\mathrm{K}} \simeq 0.2\ \mathrm{eV}. \tag{16}$$

At room temperature, the defect level appears to be significantly above mid-gap, and the hole and electron capture parameters seem to be comparable. Using next the high-N_D asymptotic dependence of $\tau_p(N_D)$ given by (1), we estimate $n_i(T_f) \simeq 5 \times 10^{15}\,\mathrm{cm}^{-3}$, or

$$T_f \simeq 330°\mathrm{C}. \tag{17}$$

If we assume that $c_p^- \sim c_n^0 \sim 10^{-8}\,\mathrm{cm}^3\text{-sec}^{-1}$, which is reasonable for an acceptor-type trap located significantly above mid-gap, then (14) and (15) imply that $N_d^0(T_f) \sim 10^{11}\,\mathrm{cm}^{-3}$. This solubility, with (9) and (17), corresponds to

$$E_a \simeq 1.4\,\mathrm{eV}. \tag{18}$$

Although we cannot, based only on these crude estimates of the defect properties, unequivocally identify the nature of the defect, it is interesting to note that these properties are characteristic of a fundamental vacancy complex in silicon. For example, (18) is approximately the activation energy of the silicon divacancy[7], and (14)–(16) are reasonable estimates of properties of the (dominant) acceptor-type level associated with the divacancy[14].

The significance of our estimate for T_f in (17) is not manifested by the absolute value of the temperature, but by the fact that it is *low* (300–400°C) relative to common silicon processing temperatures. This result thus emphasizes the importance of low-temperature processing in the fabrication of silicon devices requiring long or well-controlled carrier lifetimes.

The fundamental-limit carrier lifetime dependencies $\tau_p(N_D)$ and $\tau_n(N_A)$ derived from (7), (8), (10), (11), (12), (14), (15) and (16) are plotted in Fig. 2. Ultimate values of measured carrier lifetimes reported in [8], [11], and [12] are included in the figure to indicate the general theoretical-experimental consistency. Note that $\tau_n(N_A)$ is dramatically higher than $\tau_p(N_D = N_A)$, and that the difference increases, in accordance with Fig. 1, with increasing doping density as long as the SRH process dominates the recombination. The $\tau_n(N_A)$ data[12] for $N_A > 10^{17}\,\mathrm{cm}^{-3}$, which show a sharp decrease in τ_n with increasing N_A, signify the onset of significant band–band Auger-impact recombination[16]. Such an onset cannot be seen in the $\tau_p(N_D)$ data[8], [11] in the nondegenerate doping range because of the lower SRH $\tau_p(N_D)$ due to the higher $N_d(N_D)$.

We emphasize that the experimental data plotted in Fig. 2 represent, in most cases, only the *ultimate* carrier lifetimes measured[8], [10], [11], [12], [13], and hence are comparable to the predictions of our fundamental-limit lifetime theory. It should be noted that [13] also reports measured values of τ_n as high as 7000 μsec in 1 Ω-cm, p-type, float-zone silicon. The discrepancy between these lifetimes and the predictions of our analysis is due either to experimental error, which is likely in very-long-lifetime measurements[17], [18], or to crude assumptions in our model. This discrepancy does not however derogate our carrier lifetime theory, which predicts the observed trends—i.e. a strong $\tau_p(N_D)$ dependence showing hole lifetimes that are considerably shorter than the electron lifetimes shown by a relatively weak $\tau_n(N_A)$ dependence—and thus which encourages additional work.

DISCUSSION

The analysis described in the preceding section, which is based on measurements that have yielded the *longest*

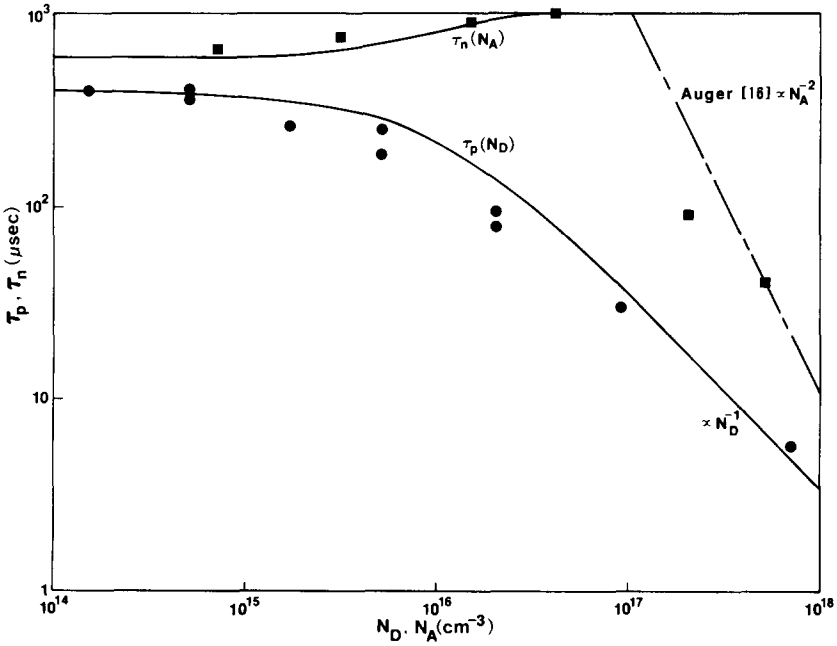

Fig. 2. Fundamental-limit minority hole [$\tau_p(N_D)$] and electron [$\tau_n(N_A)$] lifetime dependencies on doping density (continuous curves) in nondegenerate silicon as predicted by our model. The points represent the longest measured hole (circles)[8], [11] and electron (squares)[12] recombination lifetimes reported. The broken line characterizes the band–band Auger-impact electron lifetime[16], which underlies the sharp decrease in the measured $\tau_n(N_A)$ for $N_A > 10^{17}\,\mathrm{cm}^{-3}$.

recombination carrier lifetimes (→ 1 msec) reported for silicon, has implied the significance of a possibly fundamental, acceptor-type defect, e.g. a silicon divacancy. Such a defect is fundamental (unavoidable), and hence defines the fundamental limits for carrier lifetimes in nondegenerate silicon, because it is readily available for incorporation into the lattice and because its solubility is fundamentally determined by the dopant type and density.

A theoretical model for the dependence of the defect solubility on doping density, for both n-type and p-type nondegenerate silicon, was derived. This model is based on the assumptions, in accordance with the lifetime data, that the defect can be only singularly ionized and that it produces only a single important energy level. In conjunction with the SRH recombination-generation theory, our model yields theoretical dependencies of ultimate carrier lifetimes on doping density in nondegenerate silicon in which Auger-impact processes are insignificant. These results predict a strong, inverse dependence for the fundamental limit $\tau_p(N_D)$, but only a weak dependence for $\tau_n(N_A)$. Consequently in the fundamental limits, τ_n in p-type silicon will exceed τ_p in n-type silicon having the same doping density. This difference is most pronounced in low-resistivity, nondegenerate silicon.

Our model indicates the resultant defect density is dependent on an effective temperature of defect formation as defined by the silicon growth or processing. The minimum value T_f of this temperature, which results in the minimum defect density and hence the fundamental-limit lifetimes, is the temperature below which the defect is virtually immobile in the silicon lattice. Comparisons of our model predictions with the ultimate lifetime data reveal that T_f is relatively low in the range 300–400°C. Thus the final low-temperature processing steps in the fabrication of silicon devices are critical when the fundamental-limit lifetimes are being approached. These low-temperature steps must be defined so as to ensure that the fundamental defect freezes-in at T_f, and does not assume a larger solubility defined by a higher effective temperature of formation.

This conclusion suggests that annealing the silicon at T_f for a time long enough to allow the equilibrium, detailed-balance conditions (2) and (3) to obtain will minimize the density of the fundamental defect. It implies then that, with respect to this defect, T_f is an optimal annealing temperature. Annealing at temperatures above T_f will result in a higher solubility of the defect in accordance with our theory. Annealing at temperatures below T_f will be ineffective since the defect is immobile. This physical explanation is compatible with experimental determinations[18], [19] of optimal annealing temperatures that yield near-ultimate electron lifetimes in as-grown, p-type, float-zone silicon. Maximum values of τ_n generally resulted when the annealing temperature was about 450°C, which is not inconsistent with our *estimate* of T_f.

Our analysis is based on the longest carrier lifetimes measured in nondegenerate silicon, and hence pertains, as we have suggested, to a possible fundamental defect. Generally, carrier lifetimes in silicon devices are dramatically shorter than these ultimate values[3]. Shorter carrier lifetimes will result from non-optimal processing, e.g. quenching at a high temperature, that induces higher densities of fundamental defects and/or substantial densities of other, non-fundamental defects, e.g. impurities.

Iles and Soclof[20] inferred lifetimes from measurements of minority-carrier diffusion lengths in n-type and p-type silicon, and found in both types monotonic reductions in lifetime with increasing doping density. The values of lifetime they reported are much lower than those plotted in Fig. 2, and thus imply the existence of non-fundamental defects, probably heavy metals, which are common in silicon devices. Such defects, e.g. gold and copper, are generally amphoteric in silicon and tend to, because of their ionization interaction, create traps of the type opposite to the doping impurity[21]. This results in strong inverse dependencies between carrier lifetime and doping density in both n-type and p-type silicon like those in [20].

Graff and Pieper[18] investigated the dependence of carrier lifetime in silicon on heat treatment and found varying results from one silicon sample to the next because of different densities of dislocations and vacancy clusters, both of which can be avoided and hence are non-fundamental defects. They also found that in float-zone silicon the carrier lifetime-vs-annealing temperature characteristic has several maxima at different temperatures in different samples. However, there is a common maximum at about 350°C. This optimal annealing temperature is consistent with our estimation of T_f and therefore may be associated with the fundamental defect. The other maxima are probably associated with non-fundamental defects.

To achieve the fundamental-limit lifetimes that we have discussed, the non-fundamental defects must be removed or avoided by proper processing prior to the final low-temperature steps that minimize the fundamental defect density. Such processing (a) requires float-zone silicon, which is relatively free of non-fundamental defects associated with oxygen and carbon; (b) excludes high-temperature oxidation, which can create detrimental stacking faults and other dislocations and defects; and (c) includes gettering, e.g. by phosphorus diffusion, which can effectively remove certain impurities, e.g. heavy metals.

Acknowledgments—The authors thank H. J. Stein for a clarifying discussion of defects in silicon and A. Neugroschel and F. A. Lindholm for helpful comments.

REFERENCES

1. N. F. Mott, *Solid-St. Electron.* **21**, 1275–1280 (1978).
2. W. Shockley and W. T. Read, Jr., *Phys. Rev.* **87**, 835 (1952); R. N. Hall, Electron-hole recombination in germanium, *Phys. Rev.* **87**, 387 (1952).
3. B. Ross, Survey of literature on minority carrier lifetimes in silicon and related topics. *Lifetime Factors in Silicon* (ASTM STP 712). Philadelphia: American Society for Testing and Materials (1980).
4. W. Shockley and J. L. Moll, *Phys. Rev.* **119**, 1480 (1960).
5. C. S. Fuller, Defect interactions in semiconductors, *Semiconductors* (Edited by N. B. Hannay). Academic Press, New York (1975).

6. G. D. Watkins, *J. Phys. Soc. Japan.* **18** (Suppl. II), 22 (1963).
7. H. J. Stein, Defects in silicon: concepts and correlations, In *Radiation Effects in Semiconductors* (Edited by J. W. Corbett and G. D. Watkins). Gordon and Breach, New York (1971).
8. D. L. Kendall, *Conf. on the phys. and appl. of Li diffused Si*, Goddard Space Flight Center (1969); also in *Solar Cells: Outlook for Improved Efficiency.* Washington, D.C.: Space Science Board, National Academy of Science, National Research Council (1972).
9. J. G. Fossum, *Solid-St. Electron.* **19**, 269–277 (1976).
10. K. Graff and H. Pieper, Carrier lifetime measurements for process monitoring during device production, *Lifetime Factors in Silicon* (ASTM STP 712). Philidelphia: American Society for Testing and Materials (1980).
11. R. D. Nasby and J. G. Fossum, Characterization of p^+nn^+ BSF silicon concentrator solar cells, *Rec. 14th IEEE Photovoltaic Specialists Conf.* (Cat. No. 80CH1508-1) (1980); R. D. Nasby, C. M. Garner, H. T. Weaver, F. W. Sexton, and J. L. Rodriquez, Characterization of p^+nn^+ silicon concentrator solar cells, *Rec. 15th IEEE Photovoltaic Specialists Conf.* (Cat. No. 81CH1644-4) (1981).
12. H. Fischer and W. Pschunder, Impact of material and junction properties on silicon solar cell efficiency, *Rec. 11th IEEE Photovoltaic Specialists Conf.* (Cat. No. 75CH0948-OED) (1975).
13. K. Graff and H. Fischer, Carrier lifetime in silicon and its impact on solar cell characteristics, *Topics in Appl. Phys.: Solar Energy Conversion* (Edited by B. O. Seraphin), Vol. 31. Springer-Verlag, Berlin (1979).
14. L. C. Kimerling, *IEEE Trans. Nuc. Sci.* **NS-23**, 1497 (1976).
15. C. Kittel, *Intro. to Solid St. Phys.* Wiley, New York (1966).
16. J. Dziewior and W. Schmid, *Appl. Phys. Lett.* **31**, 346–348 (1977).
17. P. D. Blais and C. F. Seiler, Measurement and retention of recombination lifetime, *Lifetime Factors in Silicon* (ASTM STP 712). Philadelphia: American Society for Testing and Materials (1980).
18. K. Graff and H. Pieper, *J. Electron. Mat.* **4**, 281–298 (1975).
19. H. Fischer and W. Pschunder, Investigation of photon and thermal induced changes in silicon solar cells, *Rec. 10th IEEE Photovoltaic Specialists Conf.* (Cat. No. 73CH0801-ED) (1973).
20. P. A. Iles and S. I. Soclof, Effect of impurity doping concentration on solar cell output, *Rec. 11th IEEE Photovoltaic Specialists Conf.* (Cat. No. 75CH0948-OED) (1975).
21. R. L. Longini and R. F. Greene, *Phys. Rev.* **102**, 992–999 (1956).